

Prompting For Cost-Effective Test Ordering: A Randomized Controlled Trial

Gary R. Hollingworth¹, MD., CCFP., FCFP., Robert M. Bernstein¹, Ph.D., MD., CM., CCFP., Gary S. Viner¹, MD., CCFP., Jeffrey S. Remington², MD., CCFP., & William E. Wood², MD., CM., CCFP.

¹Medical Informatics Research Group,
Clinical Epidemiology Unit, Elisabeth Bruyere Health Centre,
²Department of Family Medicine, University of Ottawa

This randomized, controlled trial tests the efficacy of a computerized prompting system for test ordering. The system makes use of the sensitivity, specificity, positive and negative predictive values of tests. It was tested using clinical vignettes in an academic family medicine center with first and second year residents. We found that there was a 38% decrease in the numbers of tests ordered ($p < .01$) and a 12% decrease in the costs of tests ordered by using the prompting system. We suggest that when used at the point of the patient encounter, this system has the potential for promoting more appropriate test ordering and for saving considerable health care dollars.

INTRODUCTION

The costs of health care continue to soar throughout North America despite concerted efforts by governments to control them. In Ontario, the provincial government has capped the amount of money available for "fee for service" remuneration for physicians and has "clawed back" moneys from physicians if services are over-utilized in any one year. The total health care budget in Ontario is approximately \$17 billion per year with laboratory tests making up 8.9% of that amount.¹

The College of Family Physicians of Canada and other professional bodies are in the process of developing practice guidelines and quality assurance programs. The success of these initiatives will depend on the ease with which the busy practicing physician can incorporate guidelines into his/her clinical practice at the point of the patient encounter. The computerized electronic record with an integrated guidance system may be the best method of achieving this.

Although prompting systems have been in use for at least 25 years, and they have been shown to be cost-effective²⁻⁵, the utility and value of these systems in family practice has yet to be proven.

Most prompting systems tend to be too inflexible and cumbersome to be used at the point of patient care. In addition, because they are frequently rule based, they result in many suggestions being ignored⁶, and little or no learning results from their use.⁷⁻⁹ The most critical limitation of prompting systems developed to date is the fact that they do not distinguish low prevalence from high prevalence disease situations. When used in family practice, which is a low prevalence domain,¹⁰ the potential for ordering an inappropriate high number of tests exists. In settings with low prior probability, prompting for test ordering will lead to many false positive results and many unnecessary confirmatory investigations. The prompting system must therefore be flexible enough to reflect the disease prevalence milieu of the user.

PROMPTOR-FM (PRObabilistic Method of Prompting for Test ORdering in Family Medicine), is a computerized prompting system for test ordering.¹¹ This system allows the physician to rapidly display the positive and negative predictive values of a test based on one's clinical index of suspicion of disease, and to display the meaning of a positive and negative test result prior to the physician actually ordering the test. It incorporates a guidance system for pre-test likelihood of disease based on currently available evidence in the medical literature about characteristics of individual laboratory tests. The physician is able to repeat calculations of the same test and compare results with previous calculations. With this system, prompting for test ordering is uniquely tailored to each patient's situation.

We have conducted a randomized controlled trial to evaluate the effectiveness of PROMPTOR-FM in the low prevalence milieu of family practice. Our question was: Does the use of PROMPTOR-FM for test ordering change the number and costs of tests ordered by physicians?

Evidence exists which shows that a simple maneuver such as displaying the result of a previous test prior to test ordering can significantly decrease the cost of testing in Internal Medicine.⁴ In the family practice setting, it is anticipated that the physician, when faced with evidence about the accuracy of the specific test he/she is thinking of ordering, will order tests more appropriately, thus saving some valuable health care dollars.

OBJECTIVES

1. To determine the efficacy of PROMPTOR-FM in guiding physicians to cost-effective test ordering decisions.
2. To determine the potential acceptability of PROMPTOR-FM.

METHODS

This study was performed in an academic family medicine centre. First and second year family medicine residents, stratified by year of training, were randomly allocated to an experimental group and a control group.

The authors developed 5 common family practice clinical vignettes which included a short history, physical examination findings, a presumptive diagnosis and two or three questions about ordering tests in that particular situation. Table 1 contains the clinical vignettes and questions used in this study. Only in Case A is the complete scenario shown as it was presented in the study.

Table 1: Clinical Vignettes

<p>CASE A: YOUNG MALE WITH SORE THROAT 18 yo male with sore throat, first visit to you. . sore throat x 4 days, fever first day . dry cough, malaise for the last 3 days Past Hx: negative; Family Hx: negative Social Hx: university student, in residence Physical Exam: . temp 37.5 orally . throat mildly red, tonsils mildly enlarged, no pus . ears clear, clear coryza . bilateral anterior cervical adenopathy</p> <p>Based on the above, your differential diagnosis includes viral URI, Strep throat and mononucleosis. AT THIS VISIT, in addition to any other management you consider appropriate, would you order a: 1] Mono spot, 2] Throat C&S</p>
<p>CASE B: 38 yo WOMAN WITH BACK PAIN Your differential diagnosis includes back strain, disc disease and metastatic cancer. AT THIS VISIT, in addition to any other management you consider appropriate, would you order a: 1] Low back x-ray, 2] Mammogram, 3] CT scan for herniated disc.</p>

CASE C: 65 yo MAN WITH CHEST PAIN

Your differential diagnosis includes angina. AT THIS VISIT, in addition to any other management you consider appropriate, would you order an:

- 1] Exercise ECG, 2] Thallium stress test, 3] ECG.

CASE D: 30 yo WOMAN WITH ARTHRALGIAS

Your differential diagnosis includes nonseptic arthralgias, repetitive strain injury and inflammatory arthritis. AT THIS VISIT, in addition to any other management you consider appropriate would you order a:

- 1] Rheumatoid factor, 2] ANA.

CASE E: 61 yo MAN FOR ANNUAL "CHECKUP"

You wonder if screening for asymptomatic disease is in order. AT THIS VISIT, in addition to any other management you consider appropriate, would you order a:

- 1] PSA, 2] Fecal occult blood, 3] Colon flexible sigmoidoscopy.

Both groups were given exactly the same clinical vignettes.

After consenting to participate, the residents in both groups were instructed to work through each case sequentially and to answer the questions exactly as they would in their office settings. It was emphasized that their involvement was completely anonymous, they were not being evaluated and that there were no right or wrong answers.

The clinical vignettes and answer sheet were integrated with PROMPTOR-FM on computer. PROMPTOR-FM was turned on or off depending on whether a subject was in the experimental or control group. The control group residents were guided to answer the questions immediately after reading each vignette. In the experimental group, the residents were required to work through PROMPTOR-FM before making a decision on each test. Each resident took part in this study during a 1/2 hour time period during the course of a normal day of seeing patients.

After completion of the study, each resident was asked to fill in a feedback sheet and encouraged to write in any comments they felt appropriate.

RESULTS

Twenty-four residents took part in the study, twelve in each year of training, equally allocated to control and experimental groups. The time taken to complete all five clinical vignettes was <30 minutes for both groups.

For each vignette, there were two or three possible tests that were proposed as yes or no questions. Table 2 summarizes the numbers of yes and no answers

(including percentages) for each clinical vignette in each study group.

Table 2: Yes/No Answers for Each Vignette

CASE	TEST	EXPERIMENTAL GROUP		CONTROL GROUP	
		YES (%)	NO (%)	YES (%)	NO (%)
A	Mono spot	1 (8)	11 (92)	3 (25)	9 (75)
	Throat C&S	5 (42)	7 (58)	6 (50)	6 (50)
	TOTAL:	6 (25)	18 (75)	9 (37)	15 (63)
B	Back x-ray	0 (0)	12 (100)	4 (33)	8 (67)
	Mammogram	5 (42)	7 (58)	3 (25)	9 (75)
	CT scan	1 (8)	11 (92)	0 (0)	12 (100)
	TOTAL:	6 (17)	30 (83)	7 (19)	29 (81)
C	Exercise ECG	7 (58)	5 (42)	7 (58)	5 (42)
	Thallium test	3 (25)	9 (75)	1 (8)	11 (92)
	ECG	7 (58)	5 (42)	10 (83)	2 (17)
	TOTAL:	17 (47)	19 (53)	18 (50)	18 (50)
D	Rheumatoid Factor	3 (25)	9 (75)	10 (83)	2 (17)
	ANA	2 (17)	10 (83)	9 (75)	3 (25)
	TOTAL:	5 (21)	19 (79)	19 (79)	5 (21)
E	PSA	2 (17)	10 (83)	3 (25)	9 (75)
	Occult blood	1 (8)	11 (92)	3 (25)	9 (75)
	Sigmoidoscopy	0 (0)	12 (100)	1 (8)	11 (92)
	TOTAL:	3 (8)	33 (92)	7 (19)	29 (81)
Total all tests:		37 (24)	119 (76)	60 (39)	96 (61)

Overall, there were 156 yes or no answers in each group. In aggregate, the control group answered yes to 60 questions (39%) while the experimental group answered yes to 37 questions (24%). The difference of 15% between the two groups was statistically significant ($p < .01$) and this corresponded to a reduction in test ordering of 38%. The differences between the 1st and 2nd year residents was not statistically significant ($p = .327$). Table 3.

Table 3: Overall Yes/No Answers

Group	Yes	%	p	No	%	Tot
Control	60	39		96	61	156
Exp'tal	37	24		119	76	156
Diff	23	15	<.01	(23)	(15)	0
Year 1	53	34		103	66	156
Year 2	44	28		112	72	156
Diff	9	6	.327	(9)	(6)	0

The cost of each test used in this study is listed in Table 5.

Table 5: Costs of Tests in Ontario

CASE	TEST	COST (\$CAN)
A	Mono spot	7.24
	Throat C&S	17.58
B	Low back x-ray	42.23
	Mammogram	57.34
	CT scan	82.10
C	Exercise ECG	87.30
	Thallium exercise test	102.90
	ECG	20.20
D	Rheumatoid factor	7.75
	ANA	28.35
E	PSA	15.00
	Fecal occult blood	9.30
	Colon flexible sigmoidoscopy	73.65
TOTAL:		550.94

Overall, there was a 12% decrease in the cost of test ordering in the experimental group versus the control group. There was a decrease in the cost of test ordering in all cases except cases B and C. The greatest decrease occurred in case D (Table 6)

Table 6: Costs of Tests Ordered

CASE	TEST	EXPERIMENTAL GROUP	CONTROL GROUP	DIFFERENCE
		\$	\$	\$
A	Mono	7.24	21.72	14.48
	Throat C&S	87.90	105.00	17.10
	TOTAL:	95.14	126.72	31.58
B	Back x-ray	00.00	168.92	168.92
	Mammogram	286.70	172.02	(114.68)
	CT scan	82.10	00.00	(82.10)
	TOTAL:	368.80	340.94	(27.86)
C	Exercise ECG	611.10	611.10	00.00
	Thallium test	308.70	102.90	(205.80)
	ECG	141.40	202.00	60.60
	TOTAL:	1061.20	916.00	(145.2)
D	Rheumatoid Factor	23.25	77.50	54.25
	ANA	56.70	255.15	198.45
	TOTAL:	79.95	332.65	252.70
E	PSA	30.00	45.00	15.00
	Occult blood	9.30	27.90	18.60
	Sigmoidoscopy	00.00	73.65	73.65
	TOTAL:	39.30	146.55	107.25
Total for all tests:		1644.39	1862.86	218.47
Mean:		137.03	155.28	18.25

The feedback from all residents was very positive. The following characteristics were evaluated on a four point scale with 1 being best and 4 being worst; ease of use, speed, usefulness and practicality, educational, fun to use, and a final question asking if they would purchase this program. Twenty-two responded to the feedback questionnaire. The results are shown in Table 7.

Table 7: Feedback Ratings
(1 = best, 4 = worst)

Rating:	1	2	3	4
Ease of use	15	6	1	
Speed	14	6	2	
Useful/practical	7	13	2	
Educational	12	9	1	
Fun to use	13	9		
Would you buy it?	5	10	1	1

DISCUSSION

Family practice presents a vastly different patient population and disease prevalence compared to other medical specialties.¹⁰ A prompting system for test ordering must be tailor made for the low prevalence conditions of family practice. In this study, there were significant decreases in the numbers and costs

of tests ordered (38% decrease in numbers and 12% decrease in costs) when the indices of low probability of disease were incorporated into the prompting and guidance system.

One of the objectives of this study was to determine the efficacy of PROMPTER-FM in guiding the physician to make cost-effective decisions when ordering tests. A decrease in the numbers of tests ordered in a particular clinical situation may not necessarily be appropriate or cost-effective. It is important to remember that the guidance system for the pre-test likelihood used in PROMPTOR-FM is evidence based (supported by current medical literature reviews) and citations are provided for each test in the system. Current evidence for a specific situation may in fact guide the physician to order more of a particular test than past teachings or standards would have suggested.

The results pertaining to the costs of tests are less striking (overall decrease of 12%) (Table 6). There are several reasons for this. First, the pre-test likelihood guide (based on evidence in the literature) may have encouraged physicians to order more tests. Second, there is such a variability in the costs of individual tests that a large decrease in the numbers of tests in a given case may in fact lead to relatively small savings compared to a minimal decrease in the

numbers of tests ordered for a more expensive test in another case. Third, physicians may order tests more for medico-legal reasons than strictly scientific ones. Fourth, tests may be done in order to have a baseline value for future reference. When one considers the health care budget in any given jurisdiction, a 12% decrease in the costs of tests ordered amounts to significant savings.

The fact that there was no statistically significant difference in the numbers of tests ordered by the first or second year residents suggests that PROMPTOR-FM is a useful tool regardless of level of experience (Table 3).

Review of the feedback forms indicates that PROMPTOR-FM is an enjoyable, educational and easy program to use. A majority rated it as a 2 in terms of usefulness and practicality. Most of the comments on practicality indicated that the residents did not see themselves having immediate access to this program on a computer at the point of service with the patient. Only 17 residents answered the last feedback question regarding purchase of the program. The others indicated that they did not have a computer or were computer illiterate and thus could not realistically indicate an intention to purchase the program.

CONCLUSIONS

This was a small randomized controlled trial to test the efficacy of a test ordering prompting system (PROMPTOR-FM) in a family practice setting. By using this program, we found that overall, there was a 38% decrease in the numbers of tests ordered and a 12% decrease in the costs of tests ordered. The program was found to be easy, educational and enjoyable to use.

We feel that the integration of this prompting system into one's clinical practice at the point of the patient encounter, has the potential to save considerable money and at the same time, because of its evidence-based content, it will prove to be a useful educational tool. We realize that there are limitations to using vignette cases in such studies. Our next step will be to test it in community family practices with real patients.

References

1. Ontario Ministry of Health. Managing Health Care Resources. 1994-95
2. McDowell I, Newell C, Rosser WW. A Randomized Trial of Computerized Reminders for Blood Pressure Screening in Primary Care. *Med. Care* 1989;27:297-305.
3. McDowell I, Newell C, Rosser WW. Computerized Reminders To Encourage Cervical Screening in Family Practice. *J. Fam. Pract.* 1989;28:420-424.
4. Tierney WM, McDonald CJ, Martin DK, et al. Computerized Display of Past Test Results, Effect on Outpatient Testing. *Ann. Int. Med.* 1987;107(4):569-574.
5. Tierney WM, McDonald CJ, Martin DK, et al. Computer Predictions of Abnormal Test Results, Effect on Outpatient Testing. *JAMA.* 1987;259(8):1194-1198.
6. McDonald CJ, Hui SL, Smith DM, et al. Reminders to Physicians from an Introspective Computer Medical Record. A Two-Year Randomized Trial. *Ann. Intern. Med.* 1984;100:130-138.
7. McDonald CJ. Protocol-Based Computer Reminders, The Quality of Care and the Non-Perfectibility of Man. *NEJM.* 1976;295(24):1351-1355.
8. Wigton RS, Poses RM, Collins M, et al. Teaching Old Dogs New tricks: Using Cognitive Feedback to Improve Physicians' Diagnostic Judgements on Simulated Cases. *Acad Med.* 1990;65:S5-S6.
9. Poses RM, Randall D, Wigton RS. You Can Lead a Horse to Water - Improving Physicians' Knowledge of Probabilities May Not Affect their Decisions. *Med Decis Making.* 1995;15:65-75.
10. Tailoring Family Medicine Programs to Meet Community Needs. Department of Family Medicine, University of Ottawa, Can. Med. Assoc. J. 1984;131:1205-1206.
11. Bernstein RM, Hollingworth GR, Wood WE. Prompting Physicians for Cost-Effective Test Ordering in The Low Prevalence Conditions of Family Medicine. *AMIA Proceedings of the Symposium on Computer Applications in Medical Care.* 1994;18:824-828.